

INICIAR SESIÓN

NUESTROS PLANES

TODOS LOS  
CURSOS

FORMACIONES

CURSOS

PARA  
EMPRESAS

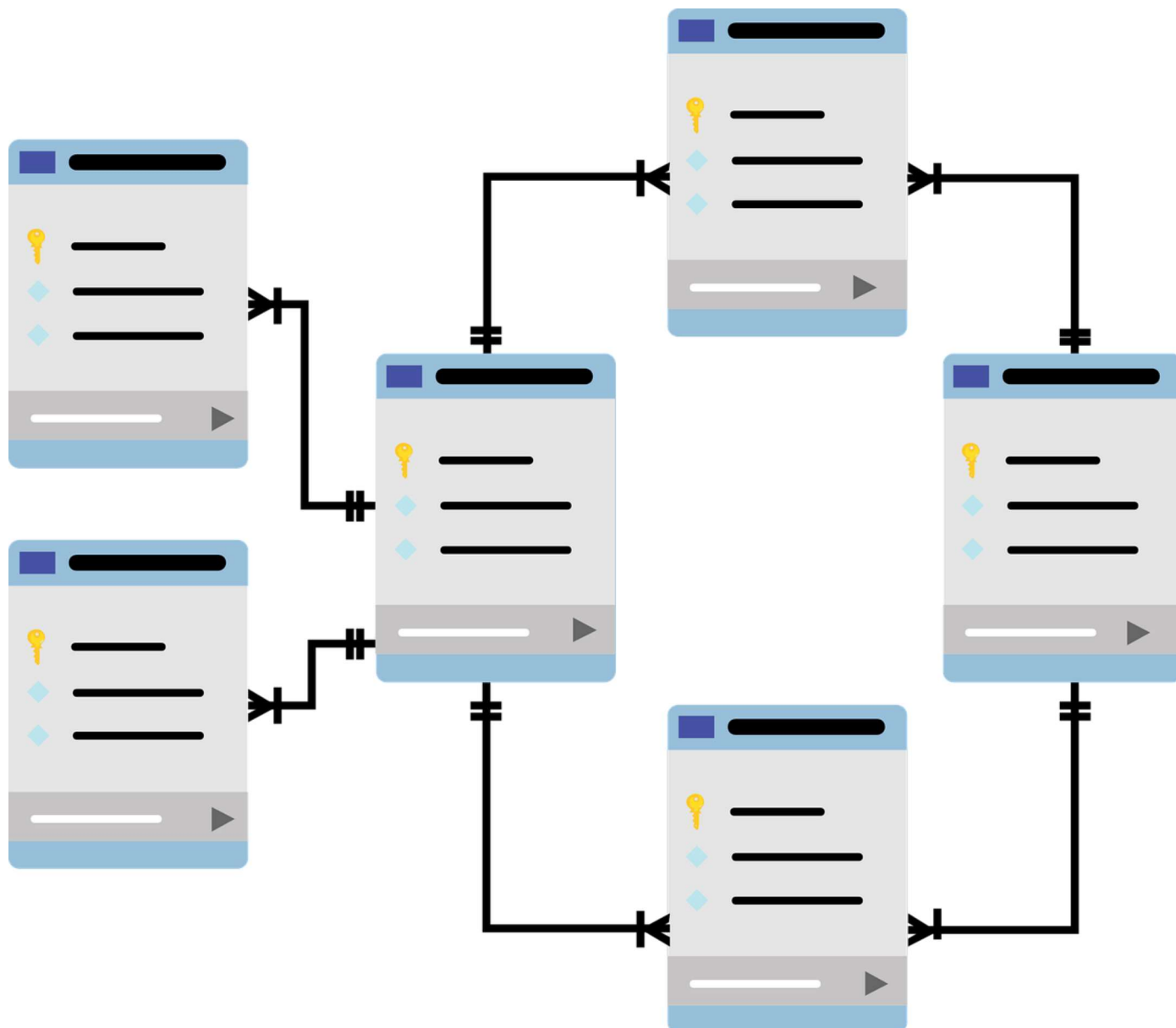
ARTÍCULOS DE TECNOLOGÍA

# Normalización en base de datos - Estructura



Luis.puig

10/10/2022



Trabajar con datos es algo cada vez más necesario, considerando nuestra realidad mayormente conectada al mundo virtual. Producimos una gran cantidad de datos todos los días y esa cantidad aumentó significativamente en las últimas décadas. Las bases de datos no están restringidas solo a las empresas, ¡las personas también tienen su propio almacenamiento! Por ejemplo, al mantener una hoja de cálculo de sus gastos personales, está almacenando datos pertinentes para su control financiero. Entonces, tener un mejor conocimiento de cómo mantener estos valores y tener buenas prácticas para manipularlos, tiene sentido.

## Normalización

La normalización es un conjunto de reglas que utilizamos sintetizadas como un proceso centrado en el modelado de la base de datos de interés. Mediante la descomposición de las relaciones presentes en la base de datos, este proceso busca anomalías, es decir,

repeticiones y redundancias entre los datos. Cuando lo encuentra, aquí es donde entra en juego el conjunto de reglas para eliminar tales anomalías y redefinir las relaciones afectadas por tal eliminación, para que todo encaje en su lugar después de los cambios. Finalmente, se enfoca en la prevención de problemas con la repetición y actualización de datos, así como el cuidado con su integridad. Este concepto fue presentado originalmente en un artículo científico publicado en IBM por el matemático [Edgar F. Codd](#), titulado "Un modelo de datos relacionales para grandes bases de datos compartidas" (1970). Codd se centraría en los valores de elementos relacionados en la base de datos, no en enlaces o agrupaciones específicas. Este modelo resultó en un proceso flexible y menos costoso para el almacenamiento y procesamiento de datos. Fue algo tan notorio que su autor ganó el [premio Turing](#) en 1981 y Forbes en 2002 marcó este modelo relacional como una de las principales innovaciones de los últimos 85 años.

## Formas Normales

Siguiendo el concepto de estandarización, tenemos estas reglas estructuradas y agrupadas en tres niveles que se utilizan para ajustar las tablas de la base de datos. Estos grupos se denominan formas normales y en este artículo se presentarán cuatro formas que se utilizan. Cada forma normal sigue requisitos y la forma anterior, es decir, se mantiene una herencia de requisitos, con excepción de la primera forma normal que no tiene la forma anterior.

### Primera Forma

De esta primera forma tratamos las repeticiones, y también nos aseguramos que los atributos se están almacenando de forma única, es decir, no hay ningún atributo con los valores en la misma fila de la tabla. Vemos la clave primaria de la tabla, y si es necesario crear otra tabla, asociamos la tabla original con la secundaria precisamente por esta clave. Vamos a crear un ejemplo con información sobre dos personas:

Tabla sin la primera forma normal.

Código	Nombre	Ubicación	Teléfono
1	Jose	Santiago	00 21 56 2 1234 7382
		Chile	00 21 56 2 1234 1234
2	Arturo	Bogotá	011-591-7-3456892
		Bolivia	011-591-7-1234656

Al principio, identificamos la clave principal, en este caso, es el atributo de código. Tenga en cuenta que entre los atributos están los valores asociados a la clave, entonces necesitamos fragmentar estos valores, y por eso es necesario crear una nueva tabla.

**Tabla con la primera forma normal.**

Código	Nombre	Ciudad	País
1	Jose	Santiago	Chile
2	Arturo	Bogotá	Bolivia

Creamos los atributos ciudad y país, porque esta información estaba en un solo atributo, ya que es más útil tener estas dos informaciones separadas, donde se relacionan.

**Nueva tabla creada a partir de la primera forma normal.**

Código	Teléfono
1	00 21 56 2 1234 7382
1	00 21 56 2 1234 1234
2	011-591-7-3456892
2	011-591-7-1234656

Esta nueva tabla la creamos para poder relacionar los teléfonos con el atributo código, que en la tabla principal es la clave primaria, y en esta se define como clave externa. Así que dejamos todos los datos definidos singularmente y, sin embargo, relacionados. En resumen, los atributos y valores posteriores a la primera forma son atómicos, es decir, son datos que no se pueden modificar ni dividir, están en su forma mínima. Además, las tablas deben contener clave primaria no nula, para efectos de identificación y relación entre datos y tablas.

## Segunda Forma

La segunda forma trabaja centrada en las posibles redundancias en las tablas, en particular, se define que los atributos de la tabla dependen enteramente de la clave primaria. Los atributos que no dependen o dependen parcialmente de ella se asignan en otra tabla, ahora con una relación clara con la clave primaria de la tabla original. Es decir, la clave primaria se convierte en clave extranjera (externa), para la nueva tabla creada. Vamos a seguir con otro ejemplo, similar a la tabla anterior.

**Tabla sin la segunda forma normal.**

Código	Nombre	Código Vuelo	Origen	Destino
1	Jose	101	Santiago	São Paulo
2	Arturo	102	Bogotá	Buenos Aires

Tenga en cuenta que los campos de origen y destino no tienen una relación directa con el campo de código, pero tienen una relación directa con el código de vuelo, ya que son información relacionada en un viaje aéreo, por ejemplo Así, podemos mover estas informaciones a una nueva tabla sin que los datos pierdan las relaciones originales.

**Nueva tabla con los datos relacionados.**

Código Vuelo	Origen	Destino
101	Santiago	São Paulo
102	Bogotá	Buenos Aires

Por lo tanto, la tabla original elimina los datos que no necesitan estar ahí, pero que siguen relacionados en una tabla secundaria.

**Tabla con la segunda forma normal.**

Código	Nombre	Código Vuelo
1	Jose	101
2	Arturo	102

Vale la pena recordar que la segunda forma normal está de acuerdo con las reglas de la primera forma normal, y así sucesivamente.

### Tercera Forma

En la tercera forma normal, trabajamos precisamente en la organización de los atributos que dependen unos de otros, pero que no son atributos clave (primarios o extranjeros). Si es necesario, cree una tabla secundaria para reestructurar la relación de dependencia entre los atributos. Estas tablas deben tener clave primaria o externa. Usemos ahora un ejemplo referente a modelos de coches:

**Tabla sin la tercera forma normal.**

Placa	Modelo	Año	Código Fabricación	Nombre de Fábrica
123X	Modelo 1	2016	1	A
456Y	Modelo 2	2012	2	B
789Z	Modelo 3	2012	3	C

Tenga en cuenta que existe una dependencia entre el atributo 'nombre de fábrica' y 'año' con el 'código de fábrica', sin embargo, estos atributos no dependen de la clave principal de la tabla que es 'placa'.

**Nueva tabla con datos de fabricación.**

Código Fabricación	Nombre de Fábrica	Año
1	A	2016
2	B	2012
3	C	2012

En este caso, creamos una nueva tabla para relacionar el nombre de la fábrica con su código, y también eliminamos las relaciones de dependencia entre atributos no clave de la tabla original.

Tabla original ahora en la tercera forma normal.

Placa	Modelo	Código Fabricación
123X	Modelo 1	1
456Y	Modelo 2	2
789Z	Modelo 3	3

Siempre es bueno recordar que para que una tabla esté en la tercera forma normal, antes de eso, debe estar definida de acuerdo con la primera y la segunda forma normal. Estas eran las tres formas normales principales. Hay una cuarta forma, no considerada la principal pero útil, que se presenta a continuación.

## Cuarta Forma

La cuarta y última forma se enfoca en eliminar dependencias multivaluadas entre los atributos de la clave, es decir, si hay más atributos (que en la clave primaria o externa) se repiten en la tabla. Si esto ocurre, generamos nuevas tablas para eliminar la redundancia y mantener las relaciones entre los atributos.

Tabla sin la cuarta forma normal.

Música	Artista	Disco
Musica 1	Artista A	Disco 1
Musica 1	Artista B	Disco 2
Musica 2	Artista A	Disco 1

Tenga en cuenta que el campo 'música' está relacionado con 'artista' y 'disco', sin embargo, el artista y la duración no pueden estar relacionados, porque sabemos que la misma canción puede estar en varios álbumes diferentes y también puede ser cantada por diferentes artistas. Por lo tanto, lo ideal es que se produzca la división de esta tabla y así eliminar las repeticiones entre los datos.

La primera tabla nueva.

Música	Disco
Musica 1	Disco 1
Musica 1	Disco 2
Musica 2	Disco 1

Y ahora el campo de la música relacionado con el campo del artista.

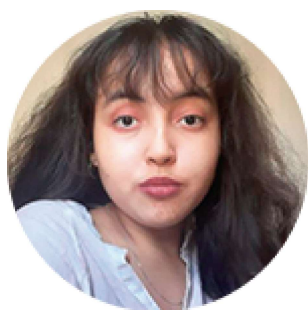
La segunda nueva tabla.

Música	Artista
Musica 1	Artista A
Musica 1	Artista B
Musica 2	Artista A

## Conclusión

La normalización es un paso importante para quien está modelando una base de datos relacional, y ciertamente redundante en una mayor eficiencia a la hora de abstraer el banco y sus atributos. Esperamos que hayas disfrutado el contenido y te animamos a que practiques este tema, para que puedas trabajar con datos de una manera más desarrollada y profesional. En la siguiente sección, presentamos algunos temas que pueden contribuir a sus estudios en la base de datos. Buenos estudios y hasta la próxima!

**Leer más** [Formación Base de Datos](#)[Artículo ¿Qué es SQL?¿Que hace un cientista de datos? Formación Ciencia de Datos](#)



**Brenda Souza** Scuba Alura LATAM. Soy estudiante de Tecnología de la Información en la Universidad Federal de Rio Grande do Norte, en Natal (Brasil). Me enfoco en lenguajes Java y Python, con áreas de interés como BackEnd, Data Science e Inteligencia Artificial. También soy desarrolladora BackEnd.





**Luis Ezequiel Puig** Soy estudiante de analisis de sistemas, formo parte del equipo de Scubas en Alura, apasionado por la tecnología y la informática. Me encanta aprender e investigar sobre nuevas tecnologías y diferentes herramientas del mundo de la programación.

ARTÍCULOS DE TECNOLOGÍA

**En Alura encontrarás variados cursos sobre . ¡Comienza ahora!**

**SEMESTRAL**

**US\$49,90**

un solo pago de US\$49,90

- ✓ 218 cursos
- ✓ Videos y actividades 100% en Español
- ✓ Certificado de participación
- ✓ Estudia las 24 horas, los 7 días de la semana
- ✓ Foro y comunidad exclusiva para resolver tus dudas



- ✓ Acceso a todo el contenido de la plataforma por 6 meses

**¡QUIERO EMPEZAR A ESTUDIAR!**

[Paga en moneda local en los siguientes países](#)

**ANUAL**

**US\$79,90**

un solo pago de US\$79,90

- ✓ 218 cursos
- ✓ Videos y actividades 100% en Español
- ✓ Certificado de participación
- ✓ Estudia las 24 horas, los 7 días de la semana
- ✓ Foro y comunidad exclusiva para resolver tus dudas
- ✓ Acceso a todo el contenido de la plataforma por 12 meses

**¡QUIERO EMPEZAR A ESTUDIAR!**

[Paga en moneda local en los siguientes países](#)

Acceso a todos  
los cursos

Estudia las 24 horas,  
dónde y cuándo quieras

Nuevos cursos  
cada semana

## NAVEGACIÓN

PLANES  
INSTRUCTORES  
BLOG  
POLÍTICA DE PRIVACIDAD  
TÉRMINOS DE USO  
SOBRE NOSOTROS  
PREGUNTAS FRECUENTES

## ¡CONTÁCTANOS!

¡QUIERO ENTRAR EN CONTACTO!

## BLOG

PROGRAMACIÓN  
FRONT END  
DATA SCIENCE  
INNOVACIÓN Y GESTIÓN  
DEVOPS

AOVS Sistemas de Informática S.A  
CNPJ 05.555.382/0001-33

## SÍGUENOS EN NUESTRAS REDES SOCIALES



## ALIADOS



En Alura somos unas de las Scale-Ups seleccionadas por Endeavor, programa de aceleración de las empresas que más crecen en el país.



Fuimos unas de las 7 startups seleccionadas por Google For Startups en participar del programa Growth Academy en 2021

POWERED BY

## CURSOS

### Cursos de Programación

Lógica de Programación | Java

### Cursos de Front End

HTML y CSS | JavaScript | React

### Cursos de Data Science

Data Science | Machine Learning | Excel | Base de Datos | Data Visualization | Estadística

### Cursos de DevOps

Docker | Linux

### Cursos de Innovación y Gestión

Productividad y Calidad de Vida | Transformación Ágil | Marketing Analytics | Liderazgo y Gestión de Equipos | Startups y Emprendimiento